

# ANALÍTICA DE RECURSOS HUMANOS: EXPLORANDO OPORTUNIDADES A PARTIR DEL *BIG DATA* Y LA PRÁCTICA DEL *HUMAN RESOURCES ANALYTICS* EXPLORING OPPORTUNITIES FROM *BIG* *DATA* AND THE PRACTICE OF *HUMAN* *RESOURCES ANALYTICS*

**David Aguado García**

Profesor Departamento de Psicología Social  
y Metodología  
Investigador Senior, Instituto de Ingeniería  
del Conocimiento  
Universidad Autónoma de Madrid  
david.aguado@uam.es

hortik kanpo gelditu, eta, pixkanaka, *Human Resources Analytics* izenarekin, hainbat praktika garatzen hasi dira, datu-analisitik abiatuta pertsonen gaineko erabakiak hartzean oinarritzen direnak. Administrazio publikoek langileen kudeaketako prozesuen optimizazioan sartu behar dute, analisiarekin sortzen diren ebidentzien bidez. Lehenik, bere profesionalei helarazi behar diete zein den gaiaren egoera, hurbilketa berri horiei dagokienez, eta, baita ere, gogoeta egin behar dute hurbilketa analitiko horien ezarpenak administrazio publikoentzat dakartzaten aukera eta zailtasunen inguruan.

*Gako-hitzak: berrikuntza giza baliabideetan, Big Data eta Giza baliabideak, Digitalizazioa, HR analytics.*

Recibido: 27/2/2018

Aceptado: 26/4/2018

**Resumen:** La explosión del *Big Data* y las tecnologías asociadas ha provocado un extraordinario cambio en la forma de hacer de las organizaciones. Este cambio afecta fundamentalmente al modo en el que estas toman decisiones para desarrollar de manera efectiva su misión. Se orientan las decisiones que tienen que ver con los sistemas de producción y provisión de servicios de forma que se maximicen sus ventajas al ser tomadas por la evidencia que provee el análisis complejo de un gran volumen de datos. Es lo que se conoce como las *Data Driven Organizations*. El campo de la Gestión de Recursos Humanos no ha sido ajeno a ello y, paulatinamente, bajo la denominación de *Human Resources Analytics*, se comienzan a desarrollar un conjunto de prácticas basadas en tomar decisiones sobre las personas a partir del análisis de datos. Las Administraciones Públicas han de incorporarse también a la optimización de sus procesos de gestión de empleados a través de las evidencias que se generan a través de la analítica. Un primer paso es acercar a sus profesionales el estado de la cuestión al respecto de estas nuevas aproximaciones y, también, reflexionar acerca de las oportunidades y dificultades que la implantación de estas aproximaciones analíticas conllevan para las Administraciones Públicas.

*Palabras clave: innovación en RRHH, Big Data y RRHH, Digitalización, HR Analytics.*

© 2018 IVAP. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Reconocimiento – NoComercial – SinObrasDerivada (by-nc-nd)



**Laburpena:** *Big Data*ren eta horri lotutako teknologien eklosioak izugarri aldatu du erakundeen jarduteko modua. Aldaketa horrek, batez ere, erakunde horiek —beren egitekoa eraginkortasunez garatzerakoan— erabakiak hartzeko moduari eragiten dio. Zerbitzuen ekoizpen- eta hornidura-sistemekin zerikusia duten erabakiak bideratzen dira, beren onurak maximizatzeko, erabaki horiek hartzen baitira datu-bolumen handi baten analisi konplexu batek ematen duen ebidentziaren ondorioz. *Data Driven Organizations* moduan ezagutzen dugun hori da. Giza Baliabideen Kudeaketaren eremua ez da

**Abstract:** The emergence of Big Data has led to an extraordinary change in the way organizations work. This change fundamentally affects the way in which they make decisions to effectively fulfil their mission. The decisions now are driven by the evidence provided by the complex analysis of a large volume of data. It is what is known as the *Data Driven Organizations*. The field of *Human Resources Management* has not been oblivious to this and, gradually, under the name of *Human Resources Analytics*, a set of practices based on making decisions about people based on data analysis begins to be developed. Public administrations must also be incorporated into the optimization of their employee management processes through the evidence generated through analytics. A first step is to put its professionals closer to the state of the art of these new approaches and, also, to reflect on the opportunities and difficulties that the implementation of these analytical approaches entail for public administrations.

*Keywords: innovation in HR, Big Data and HR, Digitalization, HR Analytics.*

## Sumario

1. Introducción.—2. *Big Data* y GRH: el campo del *Human Resources Analytics*. 2.1. Los Usos de la Analítica. 2.2. Los elementos de un proyecto *HRA*. 2.3. Las diferencias entre *Big Data* y *HRA*.—3. Nuevos datos, nueva analítica. 3.1. Nuevos recursos como fuentes de datos. 3.2. La analítica de datos. 3.3. La cuestión del criterio a predecir: el desempeño. 3.4. La necesaria mirada crítica.—4. Retos y oportunidades en la Administración Pública.—5. Conclusiones.

## 1. Introducción

Muy probablemente hablar de los avances tecnológicos que están modificando nuestro entorno y que lo van a seguir haciendo en un futuro próximo sea hacerlo del *Big Data*. Entendido como un *ecosistema* en el que confluyen diferentes tecnologías orientadas al análisis de grandes volúmenes de datos con el objetivo de proveer de buena información para la toma de decisiones, el *Big Data* se ha convertido en el elemento clave de esta, para algunos, nueva revolución industrial. Asociada y conectada también con fuerza al resurgimiento de las ideas provenientes de la Inteligencia Artificial, a los continuos esfuerzos de desarrollo de la Robotización y a la enorme eclosión de la Internet de las Cosas (Internet of Things, IoT). Esta aproximación *Big Data* no surge exclusivamente por el **volumen** masivo de datos a tratar, sino también por la **velocidad** con que se pueden generar y la **variedad** de los mismos. Son estas tres características tomadas de forma conjunta lo que se conoce como las tres V del *Big Data*. Analicemos el caso de las redes sociales: en Twitter se generan unos 6.000 tuits por segundo en un día normal (volumen y velocidad), y cada uno de ellos contiene información desestructurada como el lenguaje informal, los hashtags, o las URL (variedad).

Así las cosas, la idea es clara: existe un extraordinario volumen de datos sobre el funcionamiento de las máquinas, los sistemas y, claro, las personas: ya sean clientes, usuarios de un servicio o empleados de una organización; y, además, disponemos cada vez más de técnicas de procesamiento de esos volúmenes de datos y de análisis de los mismos. Ello nos permite enfocar el desarrollo de las diferentes actividades humanas (ya sean de negocio o de otro tipo) tomando decisiones informadas basadas en el análisis de los datos disponibles (muchos) sobre cualquier fenómeno de

nuestro interés. En el ámbito organizacional es lo que algunos han venido en denominar, las *Data Driven Organizations*.

Visto el éxito obtenido en otras disciplinas, desde la Gestión de Recursos Humanos (GRH) se comienzan a adoptar también estos planteamientos y a explorar las posibilidades que ofrece para la gestión de las personas. De hecho, la cuestión de la tecnología aplicada a la gestión de personas es un tema que aparece recurrentemente en la agenda de los principales foros sobre gestión de recursos humanos, que es señalada sistemáticamente dentro de las principales tendencias y retos a los que habrán de enfrentarse los profesionales de recursos humanos en el corto plazo y que, en definitiva, ha sido uno de los principales elementos sobre los que se está desarrollando un importante cambio en la función de recursos humanos. El fenómeno no es ni nuevo, la tecnología en sus diferentes modalidades ha estado en la base de las profundas transformaciones que en las últimas décadas se han producido en los procesos de gestión de recursos humanos; ni exclusivo, la tecnología ha sido el motor de cambio en sectores tan diferentes como el marketing, el comportamiento del consumidor, el suministro energético o el campo médico.

Como no podía ser de otra forma, si bien los desarrollos comentados anteriormente han sido realizados principalmente en el sector privado, las Administraciones Públicas se hacen eco de las ventajas de la aproximación y han comenzado a explotarla desde una posición extraordinaria: *Big Data*, va en gran medida de datos y el sector público es, a día de hoy, uno de los principales actores en lo que se refiere a la posesión de activos de información. Tal vez el ejemplo más sobresaliente sea el del desarrollo de las Smart Cities (ver por ejemplo los casos de Tokio, Nueva York o Medellín).

En este contexto, uno de los principales frenos al desarrollo del *Big Data* en las Administraciones Públicas es la identificación de casos de uso (Observatorio del Sector Público de Informática el Corte Inglés, 2016).

¿En qué contextos se puede maximizar esta aproximación? ¿Cuáles son las principales limitaciones y oportunidades en los diferentes negociados? Sin duda la GRH es uno de estos casos de uso y, tomando la experiencia desarrollada en el ámbito privado, las Administraciones Públicas pueden avanzar en el desarrollo de estrategias de gestión de personas facilitadas por la aproximación de esta amalgama de tecnologías que acompañan la aproximación *Big Data*.

En línea con lo anterior el objetivo de este artículo es plantear el estado del arte de lo que ha significado esta aproximación *Big Data* para la GRH describiendo básicamente cuáles son los impactos tecnológicos que están cambiando o atisbamos que cambiarán nuestras formas de llevar a la práctica la gestión de personas en los entornos organizacionales. Y, a partir de ahí, analizar cómo podría la Administración Pública enfocar su utilización.

## 2. *Big Data* y GRH: el campo del *Human Resources Analytics*

La relación entre la tecnología y la gestión de personas ha pasado de una primera fase en la que esta permitía liberar tiempo del profesional de recursos humanos a través de la automatización de los procesos, a una segunda fase en la que esta permite generar nuevos métodos, procesos, y formas de enfrentar la realidad que no existían anteriormente y que generan valor añadido más allá de la automatización. Así, a día de hoy el impacto de la tecnología en la gestión de personas no se basa en su capacidad para automatizar procesos existentes si no en su capacidad para generar elementos, procesos, prácticas no existentes anteriormente y que, de hecho, suponen un importante reto para los profesionales. Como decíamos anteriormente tal vez el elemento tecnológico predominante hoy en día sea el del *Big Data*. Y, en este contexto, cuando lo trasladamos directamente a la gestión de recursos humanos (a la gestión del talento, a la gestión del capital humano) aparecen términos como Analítica de Recursos Humanos, *People Analytics* o *Strategic Workforce Planning Analytics*.

Es probable que al escuchar estos términos el profesional los identifique con una especie de compendio en el que aparecen datos, indicadores, y herramientas estadísticas avanzadas. Aun no estando desencaminada la apreciación, lo cierto es que la aplicación al ámbito de los recursos humanos implica más que eso y contempla otros aspectos fundamentales como la identificación de

objetivos claros de análisis, buenos diseños experimentales que vinculen el análisis con el rendimiento de la organización, o una metodología de trabajo clara que permita involucrar en el proceso a los diferentes actores necesarios. Y, si me apuran, en último extremo, la analítica de recursos humanos trata de recuperar el modo de hacer científico en la toma de decisiones sobre la gestión de las personas en las organizaciones.

Desde esta perspectiva decimos que el Human Resources Analytics (HRA) es una metodología y un proceso integrado para proveer de evidencias, a través del análisis de datos, que sean utilizadas para mejorar la calidad de las decisiones sobre personas con el objetivo de mejorar el rendimiento a nivel individual, grupal, y/u organizacional. Se trata pues de ayudar a las organizaciones a mejorar su rendimiento alineando la «gestión y desarrollo de personas» con los objetivos del negocio. Y ello a través de una cuestión fundamental: la toma de decisiones informada.

Dos características son relevantes aquí: por un lado que las decisiones son sobre futuro y, por otro, que las decisiones buscan impactar en la misión de la organización. Y la cuestión no es baladí ya que hablar de futuro implica que los análisis tienen un fuerte componente de predicción (lo que contrasta con la mayoría de las técnicas analítico-descriptivas utilizadas en la práctica). Por otro lado, mientras los departamentos de marketing, financieros, y la mayoría de las funciones organizativas han desarrollado metodologías para generar la información que los managers necesitan para tomar decisiones estratégicas sobre el negocio, los departamentos de GRH se han centrado tradicionalmente en el análisis interno de la función sin prestar especial atención a su vínculo con el negocio. Y es tiempo de cambiar el foco: los profesionales de GRH deben comenzar a utilizar los datos no para describir y explicar el pasado sino para preguntarse y dar respuesta a las cuestiones que están a la base de la explicación de cómo los empleados contribuyen al negocio de la organización. HRA se presenta así como una herramienta imprescindible para el cambio (Harris, Craig, & Light, 2011).

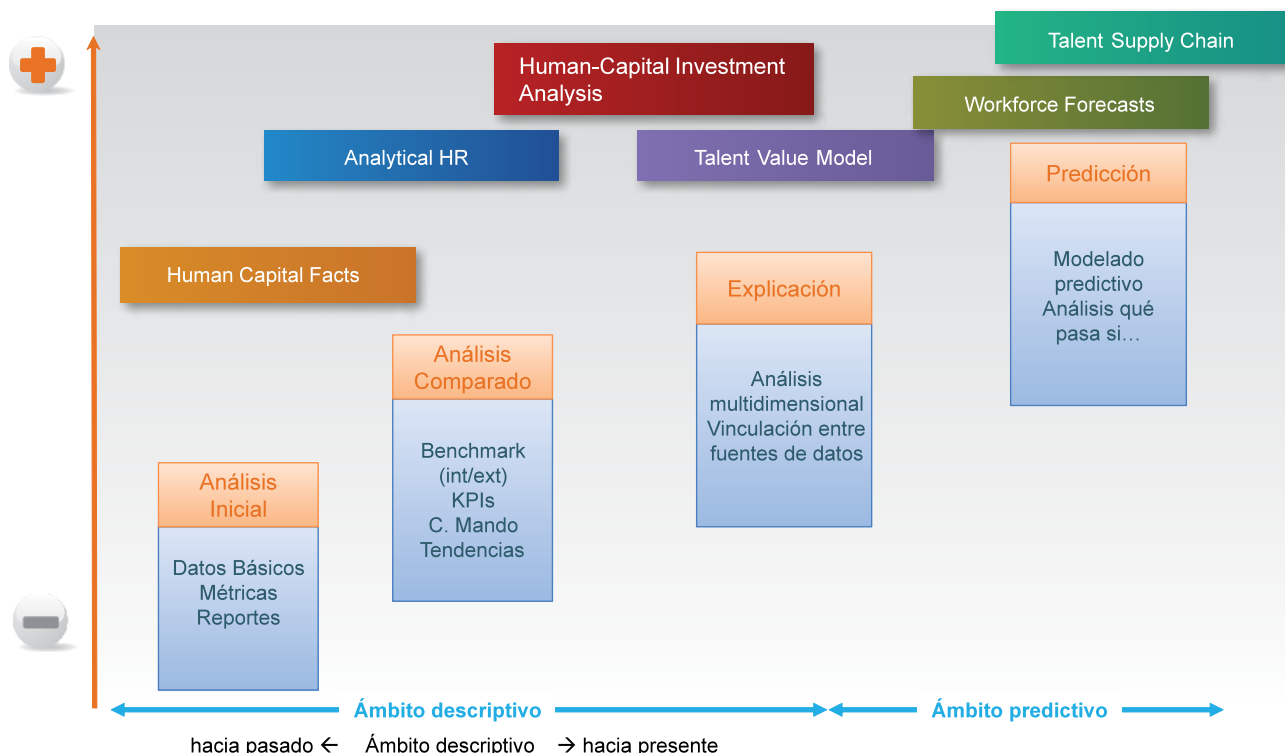
### 2.1. Los usos de la analítica

En cualquier caso, como vemos, conviene recordar que no empezamos desde la nada. De una u otra forma la mayoría de las organizaciones han llevado adelante algún tipo de utilización de la analítica de datos para tomar decisiones. Sería pretencioso por nuestra parte negar la existencia de lo que de hecho es una práctica habitual en los departamentos de recursos humanos: el análisis descriptivo de los principales indicadores de la función. Sin embargo, estos análisis están caracterizados por ser fundamentalmente descripciones del pasado. Mientras que desde



Figura 1

## Los diferentes usos de la analítica



HRA en gran medida se pretende realizar una prospectiva respecto del futuro. En la Figura 1 se muestra una representación gráfica de la madurez de las diferentes aproximaciones a la analítica de datos en GRH. Como se ve, las organizaciones varían en el grado en el que utilizan los datos de la fuerza de trabajo para tomar decisiones y, razonablemente, esas variaciones están relacionadas con el valor que aportan para la toma de decisiones.

Un primer estadio estaría constituido por aquellas organizaciones que realizan una analítica esencial: utilizan datos básicos respecto de la fuerza de trabajo y medidas *ad-hoc* respecto de un determinado ámbito. Los datos no están relacionados entre sí y se refieren a cuestiones específicas que se presentan como una descripción de lo que ha sucedido en un pasado más o menos reciente. ¿Cuántas horas de trabajo hemos facturado el último mes? ¿Cuántas horas de trabajo no están cargadas a proyectos específicos de negocio? En realidad con este tipo de analítica se trata de dar respuesta a la pregunta ¿qué ha sucedido?

Un segundo estado de desarrollo lo constituirían aquellas organizaciones que además de obtener estos da-

tos realizan con ellos un proceso de comparación respecto de lo que sucede en su sector de actuación (o bien de lo que sucede en diferentes áreas de la propia organización). Aquí hablamos de prácticas relacionadas con el *benchmarking* (interno o externo), KPIs, cuadros de mando, tendencias. En este caso la analítica permite ya una importante mejora en la calidad de las decisiones que se toman gracias, fundamentalmente, a tres hechos que acompañan esta aproximación: (a) de un lado que los datos no son tomados específicamente para la solución de un problema; (b) de otra que los datos están integrados en alguna suerte de estructura estable que favorece la sistemática de su recogida y la acumulación de estos a lo largo del tiempo; y (c) que los datos disponibles se extienden para gran parte o la totalidad de la organización. Las preguntas típicas aquí serían del estilo a las siguientes: ¿Cuál es la evolución de nuestra tasa de rotación? ¿Es mayor o menor que la de nuestra competencia? ¿Qué áreas de la organización están más satisfechas? ¿En cuáles de ellas hay menor absentismo?

Un tercer nivel de madurez se establece cuando las organizaciones se trasladan desde la comparación

hacia el ámbito de la explicación. Es entonces cuando los análisis se tornan multidimensionales y cuando, necesariamente, se vinculan y relacionan diferentes fuentes de datos. Es cuando, por ejemplo, los datos sobre el absentismo se ponen en relación con las fuentes de reclutamiento, cuando los índices de rotación se analizan a la luz de la trayectoria profesional del empleado o de, incluso, la distancia de su vivienda al centro de trabajo. En este tercer nivel la analítica se centra en la explicación de los fenómenos: ¿por qué? es la pregunta a la que se da respuesta. ¿De qué depende que unos equipos consigan mejores resultados que otros? ¿Cómo se relacionan las fuentes de reclutamiento con la rotación posterior de los comerciales?

Finalmente, el mayor nivel de desarrollo lo encontramos en aquellas organizaciones que utilizan la analítica para saltar desde la explicación de los fenómenos a la predicción acerca de cómo ocurrirán en el futuro y el pronóstico respecto de su impacto en el negocio. Aquí, necesariamente, se ponen en relación los datos referidos a la fuerza de trabajo con los datos referidos a los resultados de la organización. ¿Qué pasará si hacemos esto? ¿Qué acciones debemos tomar para mejorar estos ratios de la organización? Aquí, habitualmente, entran en juego modelos matemáticos complejos con capacidad para manejar datos de diferente tipo y aprender de las relaciones encontradas entre los datos para proporcionar escenarios de futuro. Este último nivel es lo que Boudreau & Ramstad (2007) llaman saltar el muro. Lo hacen desde la perspectiva de que transitar hacia modelos predictivos efectivos acerca del comportamiento de la fuerza de trabajo y su impacto en los objetivos de la organización implica un salto de calidad extraordinario en el valor que la analítica ofrece.

## 2.2. Los elementos de un proyecto HRA

Vistos los usos analicemos los elementos básicos que debe contemplar un proyecto de analítica para recursos humanos. Desde nuestro punto de vista cinco son los aspectos básicos: diseño de investigación adecuado, buenas preguntas, datos relevantes sobre el problema enunciado, modelos de análisis de datos adecuados y generación de conocimiento para la acción.

**Diseño de investigación.** En todo proyecto HRA la aplicación de la lógica de la investigación científica es fundamental: nos brinda el contexto de trabajo adecuado sobre el que los otros elementos del proyecto fluyen con naturalidad. La lógica de la investigación científica nos ayuda a establecer las secuencias de acción necesarias para el desarrollo del proyecto, a establecer con criterio los diferentes tipos de análisis a realizar y, en definitiva, a desarrollar una adecuada interpretación de los hallazgos encontrados.

**Preguntas.** En esta lógica de investigación la primera cuestión a plantearse en un proyecto de HRA hace referencia al objetivo del mismo. ¿A qué queremos dar respuesta? Ello implica necesariamente realizar una reflexión acerca de qué deberíamos saber respecto de nuestra fuerza de trabajo para poder tomar mejores decisiones de acción que impacten en el negocio. ¿Qué aspectos de la gestión de personas preocupan a la organización? ¿Qué resortes de GRH pueden ser utilizados para mejorar el rendimiento de la organización? ¿A qué retos debe enfrentarse la organización a través de su capital humano? Consecuentemente con la identificación de las preguntas a responder, y no menos importante, corre en paralelo la identificación (o desarrollo) de un modelo teórico que nos permita iluminar el área de acción en la que buscaremos datos para dar respuesta a las preguntas. Sin buenas preguntas es poco probable encontrar buenas respuestas. Por tanto, un elemento de enorme relevancia para el éxito de HRA es el desarrollo de preguntas adecuadas sobre qué aspectos son de interés analizar.

**Datos y métricas.** Sin datos no hay analítica. Pero datos hay muchos y, aunque los avances tecnológicos nos posibilitan cada vez más el tratamiento masivo de los mismos, desde nuestro punto de vista los datos deben ser pertinentes al problema objeto de estudio. Es decir, no valen todos los datos. Por ello una vez elaboradas las preguntas a las que dar respuesta, el siguiente paso en HRA consiste en identificar qué datos y qué métricas nos pueden dar información de interés para responderlas. Estos datos idealmente reflejarán tres tipos de cuestiones: (a) la descripción del fenómeno que queremos resolver y sobre el que versan nuestras preguntas (por ej., altas tasas de absentismo); (b) la descripción de los fenómenos relacionados con el principal que pueden influir sobre él y que en cierta medida servirían para explicarlo (por ej., barreras para asistir al trabajo); y (c) la descripción de aquellos criterios relacionados con el fenómeno en cuestión y que constituyen en cierta medida las consecuencias del mismo (por ej., baja satisfacción de los clientes). Una cuestión fundamental aquí que han señalado diferentes autores es que dependiendo del enfoque sobre el tipo de datos y métricas que se utilicen así será el valor añadido de los hallazgos obtenidos (Boudreau & Ramstad, 2002; Fitz-End, 2010). Hablamos de medidas de eficacia, eficiencia y efectividad. Las medidas de eficacia y eficiencia asociadas a los procesos tradicionales de auditoría de recursos humanos se centran en la descripción del grado en el que los procesos de gestión del talento solucionan aquello para lo que fueron desarrollados; mientras que las medidas de efectividad, mucho más próximas al HRA, tratan de reflejar el impacto de la fuerza laboral en los resultados de la organización.

**Análítica.** Una vez determinados los elementos que vamos a manipular en el análisis, llega el momento de recolectar, procesar y analizar. Para el análisis hay que considerar las fuentes de las que provendrán los datos, identificarlos, extraerlos y procesarlos para que puedan ser analizados. Evidentemente en función de lo «accesibles» que estén los datos —en términos de las capacidades informáticas que son necesarias para extraerlos—, y de la «diversidad» de fuentes sobre las que haya que trabajar, la tarea será más o menos complicada y habrá que contar con determinadas capacidades para ello. La realidad es que ni la calidad ni la organización de los datos es siempre la idónea para comenzar con el análisis, por lo que procesar los datos resulta imprescindible. El objetivo de este procesamiento previo al análisis es el de organizar los datos de forma que las diferentes técnicas analíticas puedan ser implementadas. No debemos olvidar que la calidad de los resultados obtenidos con el análisis depende de la calidad de los datos con los que se trabaja. Una vez preparados los datos llega el momento del análisis. Más allá de las técnicas estadísticas descriptivas básicas utilizadas habitualmente en los departamentos de recursos humanos, aquí se utilizan prioritariamente modelos de estadística inferencial y de análisis complejo encaminados a entender relaciones no necesariamente lineales entre los datos y a realizar predicciones con base en ellos.

**Conocimiento para la acción.** El elemento final que caracteriza HRA es la idea de «acción». El conocimiento obtenido con el análisis debe ser convertido en acciones que pueda utilizar la organización para modificar el contexto en el que opera. Esto es crítico desde nuestro punto de vista ya que el conocimiento si no es convertido en acción pierde valor añadido. Tengamos en cuenta que la velocidad a la que suceden los cambios y evolucionan los entornos en los que operan las organizaciones es tal, que la rapidez en la toma de decisiones se convierte en un elemento sumamente relevante. Así, un último escalón del HRA es implementar procesos que sigan generando automáticamente evidencias para la toma de decisiones. Se trata de «encapsular» el conocimiento hallado para que siga dando servicio a la organización en ese anhelo de la toma de decisiones en tiempo real.

### 2.3. Las diferencias entre *Big Data* y HRA

Desde el punto de vista del HRA, si bien se comparte completamente la idea de la generación de inteligencia competitiva a partir de los datos, como veremos a continuación en el terreno práctico del HRA encontramos sustanciales diferencias.

1. Con mucha frecuencia, en HRA, el volumen de datos disponibles para el análisis no es eleva-

do; más bien al contrario, disponemos de pocos datos sobre el trabajo de las personas (y sobre ellas mismas). Posiblemente existan muchos más datos sobre nuestros trabajadores en las redes sociales que en nuestros sistemas estructurados. Piensen, por ejemplo, en un coche de gama media. En él existen multitud de sensores que ofrecen datos al ordenador central para la toma de decisiones. Pero, ¿de cuántos sensores disponemos para obtener datos sobre un empleado y su trabajo para conseguir una noción más precisa de su aportación de valor?

2. Por otro lado, la variedad de estos datos es más bien escasa; en el mejor de los casos, solo existe la información estructurada contenida en los sistemas de información de recursos humanos. Sin embargo, es poco común disponer, pongamos por caso, de datos a partir de la huella digital del empleado o de su producción en espacios de creación colectiva.
3. Además, la velocidad con la que se generan esos datos dista mucho de ser la que usualmente conocemos de los sistemas automatizados en los que la producción de datos puede quedar establecida de antemano. Al contrario, la velocidad (y cantidad) de producción de datos depende de un número importante de prácticas y políticas de Recursos Humanos que desde luego no tienen una frecuencia de actualización diaria.

En nuestra opinión, y a modo de visión integradora, consideramos que el HRA como decíamos al principio es una disciplina que pretende obtener conocimiento para ayudar en la toma de decisiones a partir del análisis de las evidencias disponibles acerca de los fenómenos organizacionales. Sucede que la evolución de la tecnología, de los datos, y de las capacidades analíticas hará que, cada vez más, en los proyectos de HRA puedan ser implementadas las tecnologías del Big Data.

## 3. Nuevos datos, nueva analítica

Sea como fuere, el panorama del *Big Data* y del HRA arroja dos cambios fundamentales en el que hacer operativo de los recursos humanos: las fuentes de datos y la lógica de la analítica para la toma de decisiones.

### 3.1. Nuevos recursos como fuentes de datos

Las nuevas fuentes de datos están asociadas básicamente a contextos evaluativos. Así, hablar de tecnología aplicada para la valoración de candidatos y empleados es hacerlo de Internet y de las Social Networking Webs (SNW's). Hoy en día es común que en los procesos de reclutamiento se integren de diversas formas las SNWs (véase por ejemplo Nikolaou, 2014; Aguado, 2016) y que en el proceso de selección se utilicen tests informatizados administrados por Internet (véase Lievens y Chapman, 2009; Tippins, 2009). Tanto es así, que ya existen directrices internacionales publicadas por la International Test Commission para el desarrollo y la administración de test vía Internet (véase Bartram, 2009). Las cifras no dejan lugar a dudas: alrededor de 18,6 millones de personas de entre 16 y 74 años se conectan a Internet a diario y, a su vez, ocho de cada diez internautas de entre 18 y 55 años utiliza las redes sociales (hay más de un billón de usuarios en Facebook, más de 500 millones en Twitter y Google+ y más de 250 millones en LinkedIn), poniendo una cantidad considerable de información sobre ellos mismos a disposición de un gran número de personas.

Por otro lado, la tecnología también está permitiendo desarrollar e implementar novedosos test en los que se mejoran sus propiedades métricas (con procedimientos adaptativos basados en la Teoría de la Respuesta al Ítem), se utilizan nuevos tipos de ítems, y se generan escenarios que simulan situaciones realistas de trabajo y enriquecen la experiencia del evaluado (Reynolds y Dickter, 2010).

En definitiva esta ola tecnológica permite el desarrollo de nuevas herramientas que generan datos muy diferenciales respecto de los habituales en el mundo de los cuestionarios y, adicionalmente, nuevas capacidades para analizar esa gran cantidad de datos y obtener buen conocimiento para la toma de decisiones.

Un recorrido básico por estas nuevas formas nos pondrían en el camino de las SNWs, los mundos virtuales, los contenidos multimedia, el análisis de redes sociales, el acceso a los contenidos formativos on line y los nuevos elementos derivados del internet de las cosas.

#### 3.1.1. Las SNWs

Las SNWs aportan interesantes ventajas a los profesionales de recursos humanos: no hay costes de administración, la información la mantiene el candidato y, a menudo, la información está actualizada en tiempo real. Sin embargo, a pesar del extraordinario uso y crecimiento, existe una importante ausencia de estudios que permitan señalar a los profesionales del reclutamiento y la selección cómo utilizar con garantías

la información contenida en los perfiles SNWs de cara a desarrollar procesos selectivos de calidad (Brown & Vaughn, 2011; Kluemper, 2013; Nikolaou, 2014). ¿Cómo se producen las valoraciones de la información contenida en el SNW? ¿Qué Rol tiene la información incompleta? ¿Y la negativa? ¿Qué se está midiendo? ¿Qué correlaciones existen con otras herramientas? ¿Qué capacidad de predicción? ¿Hay impacto adverso que favorezca a determinados grupos sociales o étnicos? ¿Qué reacciones provoca en los evaluados? Son solo algunas de las preguntas que necesariamente han de ir progresivamente siendo respondidas con el objetivo de realizar una utilización de calidad de las SNWs. Desde este punto de vista, tal como señalan Roth, Bobko, Van Iddekinge, & Thatcher (2016), los SNWs debieran ser considerados y examinados como instrumentos de selección y valorados en función del estudio de su fiabilidad y validez mediante procedimientos estándar bien conocidos y aplicados en el campo del testing. Específicamente aspectos como la accesibilidad, la cuantificación estandarizada, la consistencia de la medida, su fiabilidad y validez de constructo, la capacidad de discriminación, el análisis del punto de vista del candidato/a y su utilidad son fundamentales para juzgar la bondad de los SNWs.

En esta línea diferentes estudios han mostrado como la actividad en las SNWs puede ser utilizada como evidencia de las actitudes, valores y personalidad de sus usuarios. Analizando los «me gusta» de Facebook, Markovikj, Gievska, Kosinski, & Stillwell (2013) mostraban cómo se podía predecir con un 70-90% de acierto la religión, orientación sexual y hábitos como fumar o beber, de los usuarios. Teniendo en cuenta las palabras utilizadas por los usuarios en sus publicaciones en Facebook, Schwartz et al. (2013) consiguen identificar también con precisión la personalidad de los usuarios. Similares resultados han sido encontrados por otros autores (Gosling, Augustine, Vazire, Holtzmann, & Gaddis, 2013; Quercia, Kosinski, & Stillwell, 2011; Youyou, Kosinski & Stillwell, 2015). A pesar de estos prometedores resultados, ha de señalarse que han sido obtenidos trabajando con una única SNW: Facebook. Sin embargo, no todas las SNWs cumplen con los mismos objetivos y función (Salazar-Calle & Romero-Sánchez, 2015): mientras que Facebook es una red recreativa, LinkedIn lo es profesional, siendo la red más utilizada en España en los contextos de selección de personal (Infoempleo-Adecco, 2016). Adicionalmente es necesario considerar también que no todas las redes sociales reciben el mismo grado de uso. LinkedIn (85%) y Facebook (78%) se llevan la palma mientras que My Space (13%) y Twitter (11%) tienen una involucración inferior. En cualquier caso tal como señalan las últimas directrices publicadas por la *Society for Industrial and Organizational Psychology* (SIOP, 2017) se recomienda una uti-

lización prudente por parte de los profesionales de la información contenida en estas redes.

### 3.1.2. *Mundos virtuales e In-basket informatizados*

Con diferentes grados de fidelidad, la tecnología actual nos permite desarrollar pruebas de evaluación que simulan diferentes escenarios tanto de trabajo como de juego. Los ejercicios in-basket informatizados cuentan ya con una larga tradición a la que progresivamente se van sumando el *role-playing* virtual y, más recientemente, la gamificación. Los test in-basket informatizados tratan de emular en un ordenador la dinámica de un clásico ejercicio de bandeja de los *assessment centre*. Un in-basket informatizado consiste habitualmente en una bandeja de entrada con correos electrónicos que describen las situaciones a gestionar. La información necesaria para que el evaluado pueda tomar sus decisiones se anexa a los mensajes o, alternativamente, se presenta en una página web en la que este pueda buscarla. Asimismo, puede contar con alguna herramienta para poder realizar anotaciones, presentaciones, etc. La respuesta del evaluado suele consistir en una jerarquización de las acciones que debe realizar junto con una breve descripción de cómo las gestionaría. En sentido estricto la evaluación con realidad virtual utiliza ejercicios de evaluación en los que tanto la presentación de información al evaluado como su respuesta se realizan a través de dispositivos de realidad virtual. Sin embargo, lo más común es que el ejercicio de evaluación tome la forma de un escenario simulado en un mundo virtual en 3D como Second Life, Wonderland o Croquet. Cuando en ese mundo virtual se introducen elementos que reproducen exactamente objetos reales como edificios, lugares de trabajo, etc., se habla de realidad aumentada. En este tipo de ejercicios el evaluado adopta el papel de un avatar y, a partir de ahí, interactúa con otros avatares y objetos. Lo que caracteriza estos escenarios es que en ellos el evaluado realiza una navegación libre y utiliza diferentes objetos 3D a los que se asocian diversas funcionalidades. Cuando el evaluado interactúa con los elementos del mundo virtual en un diseño que responde a los parámetros básicos de un juego (con capacidad para divertir y entretener), hablamos de gamificación. Además de las ventajas obvias de un entorno de evaluación en el que se simula un determinado contexto laboral (o lúdico), los 3D Virtual Worlds proveen dos importantes elementos de cara al diseño de pruebas de evaluación: entornos en los que diferentes evaluados interactúan entre sí (de manera síncrona o asíncrona) y la posibilidad de registrar de forma no intrusiva tanto la interacción con otros evaluados y objetos presentes en la simulación como el tiempo y lugar en el que se producen.

### 3.1.3. *Test Multimedia de Juicio Situacional (TMJS) y Test Objetivos de Personalidad (TOP)*

Sobre la base de los Test de Juicio Situacional, en los TMJS se incluye la tecnología a través de diferentes elementos: las situaciones pueden presentarse con una grabación en vídeo y las respuestas del evaluado pueden obtenerse en un formato escrito y abierto o en una grabación en vídeo. La presentación en vídeo de las situaciones deriva en dos aspectos destacables: el evaluado otorga una mayor validez percibida al test, y la demanda cognitiva para entender la situación es menor que en un formato escrito, lo que incrementa el número de perfiles profesionales que pueden ser evaluados. La ventaja del formato de respuesta abierta (ya sea un vídeo o un texto escrito) frente al formato de elección de alternativas es clara: el evaluado no ha de «elegir» o «reconocer» la respuesta adecuada, sino que debe «crearla», con lo que es más plausible obtener una expresión directa y espontánea de la competencia del evaluado. Por otro lado, mientras que en un test de autoinforme se aprecia el comportamiento del evaluado a partir de las verbalizaciones sobre su propio comportamiento, en un TOP se hace a partir de su comportamiento en un contexto planteado para observar un determinado tipo de conductas. Visualmente, el evaluado se enfrenta a una tarea interactiva en la que ha de «hacer» cosas. Esta tarea está especialmente diseñada para conseguir que manifieste su tendencia natural de comportamiento respecto de las variables a evaluar. Una de las ventajas más sobresalientes de este tipo de test es que se aprecia el comportamiento real del evaluado en un contexto específico y no una apreciación verbal acerca de su forma de sentir, pensar o actuar. Con ello se consigue una mayor resistencia a los sesgos relacionados con la deseabilidad social. Asimismo, cuanto más se acerque el comportamiento observado en el test al que debería desarrollar en el puesto de trabajo, mayor es la probabilidad de que el comportamiento observado en la prueba sea similar al que desarrollará el evaluado en el puesto de trabajo, de ahí la importancia de este tipo de ejercicios en la predicción del desempeño futuro del evaluado. Un elemento importante es que el soporte tecnológico para las anteriores herramientas posibilita la captura de una cantidad extraordinaria de datos de los evaluados: no solo se consideran ya las respuestas correctas o incorrectas sino que, más allá, cobra interés el proceso que siguen las personas para interactuar con la prueba de evaluación. Ello constituye un enorme reto de modelado y de interpretación del comportamiento a partir de los datos generados. Si a ello se le suma la posibilidad de que el movimiento del evaluado e incluso sus expresiones faciales sean grabadas mientras interactúa con los ejercicios a través de dispositivos como la Kinect© el coctel es completo. Iniciativas en



este sentido ya han sido desarrolladas y prometen ser un campo fecundo (Delgado-Gómez et al., 2015; Rubio, Aguado, Delgado-Gómez, Márquez, & Cernada, 2017).

### 3.1.4. Análisis de Redes Organizativas y Sociales

Cada vez más el análisis de las interacciones entre los individuos se convierte en una gran fuente de datos. Mediante este tipo de datos se obtiene información acerca de la «red» en la que una persona se sitúa (Haya y García, 2018). En los últimos años ha cobrado fuerza el Análisis de Redes Sociales como aproximación útil en la descripción de los fenómenos organizacionales. Como disciplina de estudio surge en la primera mitad del siglo XX, muchos lo conocimos a través del socio-grama, y el auge de las plataformas de interacción social ha hecho que se produzca un verdadero reverdecimiento de este tipo de análisis con el desarrollo de modelos y técnicas matemáticas de gran potencia. Mediante esta aproximación se trata de reflejar una determinada realidad organizacional a través del estudio de las relaciones que existen entre los diferentes miembros de una organización o de un grupo. En el análisis se contemplan dos tipos de elementos: los nodos (las personas) y los enlaces (las relaciones entre los nodos). La idea se puede observar en un ejemplo sencillo a escala muy pequeña. Tenemos un grupo de 4 personas en un determinado departamento e indicamos a cada una de ellas que elija a quién/quienes del resto acudiría para solucionar un problema técnico. Mediante este tipo de datos se representan las relaciones encontradas con esa sencilla pregunta. Lo relevante del asunto está en que, a menudo, las relaciones entre los miembros de una determinada organización expresadas a nivel formal nada tienen que ver con la dinámica informal que da vida al desempeño de los empleados. Mediante este tipo de datos podemos acceder a esta red informal tanto en pequeños como en grandes grupos.

### 3.1.5. La adaptación de los procesos on line de formación

Sin duda la innovación tecnológica más relevante en el campo relacionado con la formación y el desarrollo nos ha llegado de la mano de la formación on-line. La idea de la formación a distancia cristaliza con la era de Internet en el elearning generando nuevos modelos de enseñanza-aprendizaje que, de hecho, han impactado en gran medida en cómo se desarrollan los procesos formativos en las organizaciones. En este contexto, el profesional de la gestión de recursos humanos ha de considerar necesariamente el hecho de que un nuevo modelo de formación requiere de la redefinición de algunos de los parámetros básicos sobre los que ha descansado la formación tradicional

durante años. Este nuevo modelo requiere de: (a) un nuevo concepto de formador: el tutor, con unas competencias que son diferenciales respecto de las del formador en entornos presenciales y que tienden no tanto a una alta capacitación técnica sino más bien a un desarrollo como facilitador del proceso de aprendizaje individual que realizan los alumnos; (b) también es necesario un nuevo alumno con responsabilidad sobre su propio proceso de aprendizaje y con motivación para llevarlo a cabo en el marco de su plan individual de desarrollo, lo que entronca claramente con la visión de un empleado que puede acceder a la mayoría de los contenidos formativos que desea y que es el principal responsable de su crecimiento profesional; (c) requiere también de una visión diferente del aprendizaje: más basada en los principios del aprendizaje colaborativo, el *learning by doing*, y el aprendizaje constructivo, y menos reposada sobre la tradicional lección magistral; y (d) la concepción de un nuevo espacio para el proceso de formación como es el aula virtual, aula que es ubicua, accesible a demanda del alumno y que rompe con la necesidad de sincronía espacio-temporal entre los alumnos, el profesor, y el contenido. En definitiva un contexto formativo en el que el contenido está encapsulado y no es propiedad del formador, en el que los alumnos acceden a éste cuando quieren, y en donde el aprendizaje se produce cuando el alumno interactúa con dicho contenido, con los otros alumnos y con el facilitador del proceso de aprendizaje.

Sobre esta base los últimos años han venido a sorprendernos con las actualizaciones del modelo para incorporar la eclosión de las redes sociales y de los principios de la web semántica. Downes (2008) ha acuñado el término elearning 2.0 para hacer referencia a la orientación del aprendizaje hacia un entorno más social y dinámico en el que las redes sociales juegan un papel principal; y algunos autores como Bersin (2004) hablan de un elearning 3.0. soportado sobre entornos colaborativos en los que la presencia de un tutor o profesor que dirija el aprendizaje es sustituido por los modelos y aproximaciones automáticas derivadas de la ciencia computacional y la web semántica. Sin embargo, a nuestro juicio, el desarrollo más relevante producido en los últimos años tiene más que ver con un cambio estratégico en el paradigma que con un cambio estrictamente tecnológico. Nos estamos refiriendo a la formación on line masiva en abierto (Massive Open Online Course, MOOC). Más relevante que las características de cada curso lo es el hecho del planteamiento global que se realiza con esta aproximación: cualquier persona puede acceder gratuitamente a cursos diseñados por las mejores universidades en una variedad extraordinaria de disciplinas. En una revisión realizada *ex profeso* para este artículo encontramos más de 2.000 cursos disponibles en las principales plataformas universitarias que ofre-

cen MOOCs-Coursera, edX, Miríada y Future Learn. Sin duda esta es una buena base para que el trabajador planifique el desarrollo de su carrera y acceda a los elementos formativos que necesita para ello. Ello implica un cambio extraordinario en cómo el profesional de la gestión de recursos humanos se integra en el proceso de desarrollo de carrera del empleado.

### 3.1.6. Las innovaciones por venir: *IoT*s, *APP*s y *Wearables*

La idea que queremos reflejar bajo este epígrafe hace referencia a que muchos de los objetos cotidianos con los que interactuamos ya generan datos y están interconectados con otros a través de internet y que, de manera adicional, existe una gran cantidad de aplicaciones y dispositivos gestionados a través de los móviles que nos permiten interactuar entre nosotros y con estos otros sistemas. Un ejemplo para entenderlo: las salas de reuniones de la mayoría de las grandes organizaciones se encuentran «digitalizadas» su acceso se realiza desde una aplicación, se sabe quién va a utilizar la sala en un determinado momento, cuando entran y cuando salen de la sala, etc. Pues bien, cuando los asistentes a una reunión salen de la misma, el «sistema» puede enviar un mensaje a la aplicación móvil de los asistentes para que valoren la calidad de la reunión. ¿Han conseguido sus objetivos? ¿El ambiente ha sido cordial? Y un sinnúmero de posibles cuestiones. Los datos generados así responden al criterio de obtención en tiempo real y su almacenamiento y análisis posterior redunda en una buena información para la toma de decisiones.

## 3.2. La analítica de datos

La segunda cuestión relevante en esta aproximación HRA es que se pretende utilizar más allá de las técnicas multivariadas de la estadística convencional, el conjunto de técnicas analíticas desarrolladas en el campo de la ingeniería algorítmica y que actualmente etiquetamos con el término genérico de técnicas de reconocimiento de patrones como el *Machine Learning*, *Deep Learning* y términos similares. Un patrón puede definirse como una determinada regularidad existente en una fuente de datos (Shawe-Taylor & Cristianini, 2004), se han desarrollado diferentes técnicas estadísticas con un fuerte componente computacional que se han aplicado con indudable éxito a distintos problemas existentes en diferentes campos de investigación: la identificación de regiones que contienen caras en imágenes digitales (Pavani, Delgado, & Frangi, 2010), la visión por ordenador (Cheng, Zheng, & Qin, 2005), las finanzas (Tay & Cao, 2001), el campo de la genética (Guyon, Weston, Barnhill, & Vapnik, 2002) o la clasificación de textos (Tong & Koller, 2002) por poner solo algunos ejemplos. Sin

embargo, a pesar de la amplia utilización que han tenido estas técnicas en otros campos, es realmente escasa su utilización en ámbitos HR. Probablemente dos son los motivos principales por los que estas técnicas han tenido tan limitado uso en nuestro campo: de un lado la complejidad de las técnicas y, de otro, su consideración de *caja negra*. Respecto del primero la cuestión parece clara: la utilización de estas técnicas requiere de profesionales con una formación matemático-estadística que se aleja con mucho del perfil profesional que se involucra en los procesos HR (por mucho incluso que estos tengan una fuerte componente analítica). El segundo motivo refiere a una cuestión teórica que, a mi modo de ver, tiene gran importancia: aunque estas técnicas mejoran considerablemente la precisión de las clasificaciones, el modo en el que los predictores son tenidos en cuenta para producir una determinada clasificación o predicción está oculto para el investigador (de ahí la denominación de *caja negra*). Ello sin duda genera un conflicto claro ya que podemos predecir con cierta precisión la ocurrencia de determinados fenómenos pero no podemos explicar por qué. En el ámbito del HRA esto tiene vital importancia ya que las decisiones que se toman a partir de las predicciones afectan a las personas. En este contexto la comunicación clara del soporte utilizado para tomar determinadas decisiones se impone como una cuestión fundamental para la credibilidad de HR y, en este sentido, las *cajas negras* no son precisamente un aliado para la transparencia.

En cualquier caso la precisión de estas técnicas suele ser superior a las técnicas estadísticas multivariadas (Delgado-Gómez, Aguado, López-Castromán, Santacruz y Artés-Rodríguez, 2011) y se impone su utilización de forma que pueda limitarse la incidencia del efecto de *caja negra*. Desde una mirada sustantiva la mejora de la capacidad predictiva que permiten estas técnicas radica fundamentalmente en su carácter no lineal y por ello son más flexibles para modelar una realidad que, en la mayoría de los casos no tiene por qué responder a una configuración de relaciones lineales.

Con estas bases diferentes trabajos aplicados han mostrado su eficacia. Por ejemplo, Chien y Chen (2008) utilizando la edad, el género, el estado civil, el nivel de educación, la formación académica y la experiencia laboral de los candidatos para trabajar en una organización del sector industrial, implementan un sistema basado en árboles de decisión, que permite a los gestores de Recursos Humanos predecir el rendimiento futuro del candidato, la posibilidades de fidelizarlo y sus razones para abandonar la compañía. También en el terreno de la retención del talento se han obtenido resultados satisfactorios (Quinn, Rycraft, & Schoech, 2002; Tzeng, Hsieh, & Lin, 2004) utilizando redes neuronales y Support Vector Machine consiguiéndose pre-

decir el número de empleados y supervisores que se marcharían de la compañía en un tiempo reducido con porcentajes de acierto de entre el 80% y el 89%. Delgado-Gómez et al., 2011) utilizan también este tipo de técnicas para la predicción del rendimiento consiguiendo niveles de precisión cercanos al 90%. En la misma línea, Cho y Ngai (2003) desarrollan el *Intelligent Agent Selection Assistant for Insurance* en el que implementan un conjunto de algoritmos sobre el sistema de información de una gran firma de seguros de Hong Kong, que ofrece al profesional GRH una predicción certera acerca de cuánto tiempo colaborará el candidato con la compañía, qué nivel de ventas alcanzará e, incluso, qué volumen de incentivos obtendrá.

A este respecto, una aplicación interesante de la planificación de Recursos Humanos es la aportada por Yang, Lin., Chen, & Shi, (2009). Utilizando minería de datos en la previsión de la fuerza de trabajo en las unidades de emergencia del sistema hospitalario de Taiwán, consigue una optimización del proceso y un ahorro de costes en el personal médico cercano al 37%.

En definitiva, las técnicas y procedimientos derivados del Big Data están suficientemente maduros como para integrarse en sistemas de selección y asignación de personal que hagan hincapié en el tratamiento de los diferentes *inputs* generados en los procesos selectivos y que generen información relevante para tomar decisiones de contratación en términos de predicción del rendimiento futuro de los candidatos (en puestos específicos o en familias de puestos laborales) y del grado de retención de los mismos.

Un campo de enorme expansión para la analítica está siendo el de las MOOCS y el eLearning. Grandes volúmenes de estudiantes y grandes volúmenes de datos recogidos de las interacciones de cada uno de ellos con los sistemas. Ello ha hecho que a través de la analítica se hallan desarrollado interesantes aproximaciones (Romero y Ventura, 2010). Procedimientos para personalizar el proceso de aprendizaje, recomendando a los alumnos las actividades que puedan aportar más valor a su aprendizaje en el ritmo adecuado para lograrlo; ofreciendo a los alumnos *feedback* automatizado específico y de calidad respecto de sus puntos fuertes y débiles, así como estrategias para afrontarlos; construyendo modelos analíticos de la interacción de los alumnos con los contenidos formativos para mejorar los diseños instruccionales; delimitando las mejores formas de organizar los contenidos y los recursos formativos específicos (por ejemplo, los contenidos y recursos de una universidad corporativa) en función de su aprovechamiento futuro en términos de desempeño; e identificando a aquellos estudiantes con más posibilidades de abandonar y determinando las posibles causas y acciones de intervención.

### 3.3. La cuestión del criterio a predecir: el desempeño

Hablar de Big Data y de Analítica es hacerlo en gran medida de predicción. Y, en este contexto, el desempeño en su concepción multidimensional (Rotundo & Sackett, 2002) es el criterio a predecir por excelencia. Los avances en la analítica y la evolución en las nuevas fuentes de datos nos permiten realizar una nueva mirada al complejo mundo del criterio. Por ejemplo los últimos estudios relativos a la distribución del desempeño organizacional ponen de manifiesto que lejos de responder a una curva en forma de campana más bien responden a lo que denominamos una curva de *long tail* (Aguinis & Bradley, 2015). En este tipo de distribuciones se asume que existe una proporción alta de puntuaciones muy superiores a la media y que, además, la mayoría de las puntuaciones caen por debajo de ella. Ello tiene implicaciones extraordinarias en el análisis del desempeño como por ejemplo la existencia de un mayor número de *star performers* que los pronosticados bajo los anteriores modelos. Y, a la vez, un volumen muy amplio de trabajadores por debajo de la media. Las prácticas de gestión así se ven necesariamente modificadas. Otro aspecto de indudable interés está en la apreciación de las contribuciones de los trabajadores: estas contribuciones pueden ser estimadas a partir cada vez más de diferentes fuentes de datos obtenidos, además, de manera contingente a cuando se produce la aportación de valor. En cierto sentido la *sensorización* progresiva de la actividad laboral nos encamina a la obtención de datos en (casi) tiempo real acerca de los comportamientos y resultados de los trabajadores.

### 3.4. La necesaria mirada crítica

Datos, analítica y personas. Cojamos todos los datos que necesitemos, apliquemos los análisis que correspondan y tomemos decisiones basadas en la evidencia. Fácil ¿no? Sin embargo, ¿podemos utilizar cualquier tipo de dato? ¿Aun pudiendo, debemos? ¿Podemos tomar cualquier decisión aun con un margen de error pequeño? ¿Podemos etiquetar a una persona tomando como base un modelo predictivo de carácter probabilístico? ¿Qué sucede en este caso con el fenómeno de la profecía auto-cumplida, de la probabilidad de error? Es claro que los profesionales de recursos humanos necesitan ponerse de acuerdo acerca de un conjunto de buenas prácticas que puedan servir de estándar de acción de forma transversal en cualquier organización. Para ello conviene aclarar algunas cuestiones relacionadas con los datos, con el procedimiento y con los resultados.

Respecto de los datos lo primero a resaltar es que recientemente disponemos de un marco regulatorio de

obligado cumplimiento al que hemos de recurrir: el reciente Reglamento General de Protección de Datos dictado por la Unión Europea efectivo a partir del 25 de mayo de 2018. En las regulaciones al respecto que deberán ser atendidas se establecen las figuras de propiedad y protección de los datos.

Independientemente de la cualidad de los datos, de la propiedad de los mismos y de los niveles de seguridad aplicados, hay algunos mecanismos a aplicar y que identificamos como una buena práctica que redundará en la transparencia del proceso y en la seguridad de los empleados.

El primero de ellos es la pseudonimización o descaracterización de los datos a emplear en los análisis. En un archivo típico para ser analizado los datos estarán asociados a empleados concretos y los empleados estarán identificados con algún código biunívoco (como por ejemplo el DNI o el código de empleado). La buena práctica consiste en eliminar ese código que permite identificar a un empleado en concreto y sustituirlo por algún tipo de código también biunívoco pero que no nos permita llegar al empleado específico al que se refieren los datos. Cuando esta sustitución es irreversible (no podemos volver atrás de ninguna forma) hablamos de descaracterización; cuando sí podemos (por ejemplo porque hemos generado el nuevo código con algún algoritmo o porque podemos guardar en un archivo diferente la relación código original con código del análisis) hablamos de pseudonimización.

El segundo de ellos tiene que ver con un procedimiento que ha cobrado fuerza en la investigación académica en los últimos años y que debe ser trasladada con claridad a los proyectos de analítica para recursos humanos: el consentimiento informado. Esto es una declaración expresa de los empleados para los que analizaremos sus datos de que conocen los objetivos del proyecto, el procedimiento y los posibles usos que pueden darse a los resultados, y consienten expresa y voluntariamente en que los datos referidos a su persona sean utilizados.

Finalmente, una cuestión que nos encontramos con frecuencia en los proyectos de analítica hace referencia al nivel en el que se devuelven los resultados. No es lo mismo hacer un pronóstico del volumen de absentismo que tendrá un determinado departamento de la organización a dos meses vista que clasificar a los empleados de ese departamento atribuyéndoles una probabilidad de tener unas determinadas ausencias en el mismo periodo de tiempo. Lo primero es ofrecer resultados globales que permiten a la organización desarrollar un conjunto de prácticas para limitar las ausencias y/o paliar sus efectos previendo que sucederán. Lo segundo ofrece la posibilidad de bajar al caso individual y poner el foco en un empleado específico

o en grupo de empleados determinados en lo que se conoce como perfilado. La situación es que si por un lado son indudables las ventajas que el perfilado puede tener (relacionadas con la detección precoz de los acontecimientos y con la detección de grupos sobre los que focalizar la acción), también existen importantes desventajas relacionadas con la generación de estereotipos, prejuicios y discriminación (Brown, 2010).

## 4. Retos y oportunidades en la Administración Pública

En líneas generales podríamos decir que a través de la utilización de la analítica y las aproximaciones *Big Data* las Administraciones Públicas estarían en una buena disposición tanto para desarrollar mejores políticas públicas a través de un ajuste más certero respecto de las demandas de los ciudadanos, como para tomar decisiones más eficientes en el corto plazo en el ámbito del mantenimiento del servicio a los ciudadanos (por ejemplo, relacionados con el suministro de los diferentes servicios energéticos o la distribución de las redes de transporte). El ámbito de los recursos humanos también ha de ser incluido en estos *casos de uso*. Los servicios prestados a los ciudadanos se desarrollan a partir de la interacción de estos con los empleados públicos a diferente nivel. Por tanto, la acción de la analítica en este ámbito se torna en imprescindible para una mejora de los servicios públicos prestados por las diferentes administraciones. Conectado con ello está la visión de que es el ciudadano el que está en la visión de la gestión de recursos humanos. En un reciente artículo, Ulrich & Dulebohn (2015), realizando un análisis de la evolución de la función de recursos humanos establecen que la nueva «ola» en la que se verá inmersa esta gestión para generar valor añadido es lo que denominan la aproximación de *outside/inside*. Desde este punto de vista los departamentos de recursos humanos no crean valor únicamente a través del servicio a los empleados o el rediseño de las prácticas de gestión haciéndolas más eficientes. Si no que, más bien, el enfoque consiste en crear valor a partir del aseguramiento de que los servicios de recursos humanos ofrecidos dentro de las organizaciones se deben alinear con las expectativas que sobre esta se tiene fuera de ella. Máxime, añadimos nosotros, en un contexto como el de los empleados públicos y el servicio al ciudadano. Las prácticas de recursos humanos por tanto se re-diseñan para que generen valor fuera de la organización. Se introduce así un eje externo con el que hay que conectar



la gestión que redundará en una mayor ventaja competitiva de la organización. Es precisamente en este contexto en el que la capacidad de los profesionales de recursos humanos para añadir valor en el futuro requerirá de inversión en programas y proyectos específicos de analítica. Ello conllevará necesariamente el desarrollo de nuevas habilidades y competencias en los profesionales de recursos humanos para comprender como utilizar la analítica para tomar decisiones y qué tipo de análisis son necesarios y cuándo.

Sin embargo, la particularidad de las Administraciones Públicas hace que lo anterior no pueda trasladarse sin más. Como señalan Pin, García y Peiró-Barral (2012) las Administraciones Públicas se enfrentan

al reto del desarrollo de estrategias de motivación y gestión eficiente del desempeño en un contexto en el que la limitación presupuestaria es clave. Por otro lado, la fuerte extensión del modelo burocrático de gestión limita la flexibilidad tantas veces necesaria para aprovechar este contexto de grandes volúmenes de datos y altas capacidades analíticas. En la Figura 2 hemos dibujado el análisis DAFO de la cuestión. En el reflejamos lo que desde nuestro conocimiento pueden constituir las principales fortalezas y debilidades de las Administraciones Públicas para la aplicación de este conjunto de tecnologías, y, en su visión externa, las amenazas y oportunidades que el contexto actual ofrece.

Figura 2

## DAFO sobre la aplicación de la analítica de recursos humanos en el sector público

### Debilidades

- Limitada actitud favorable ante el cambio
- Lentitud a la hora de implantar nuevos procesos y/o formas de hacer
- Ausencia de dirección estratégica compartida
- Ausencia de mecanismos eficaces abajo-arriba
- Limitada capacidad de riesgo

### Amenazas

- Data infoxización
- Ausencia de persistencia en el desarrollo de modelos y teorías
- Incapacidad para desarrollar un marco de acuerdo de aplicación
- Baja capacitación de los profesionales de recursos humanos

### Fortalezas

- Gran volumen de datos
- Fuerte control de procesos y recursos
- Necesidad compartida de mejora
- Vocación de servicio - significado del trabajo
- Desarrollo de nuevas capacidades

### Oportunidades

- Capacidades tecnológicas adaptadas
- Desarrollo de buenas prácticas y casos de uso
- Involucración de las diferentes administraciones y a diferente nivel

**Debilidades.** En cuanto a las principales debilidades de la función pública para implantar con éxito proyectos de analítica para recursos humanos encontramos una limitada actitud favorable ante el cambio; lentitud a la hora de implantar nuevos procesos y/o formas de trabajo; ausencia de un programa de dirección estratégica fuertemente compartida por los diferentes niveles organizativos; ausencia de mecanismos eficaces de comunicación abajo-arriba y una limitada aceptación del riesgo. Como se ve, estas debilidades están relacionadas fuertemente con la flexibilidad de la Administración Pública para implantar y desarrollar cambios. También, con la necesaria visión en la que se integra, en la formulación de la estrategia, la información que proviene de los puntos de contacto que el empleado público mantiene con el ciudadano. Ello

deriva en una dificultad para desarrollar prácticas y políticas basadas en la evidencia que proviene de los datos generados en estos niveles. Por otro lado, aunque las Administraciones Públicas están en una buena posición para desarrollar esfuerzos sostenidos en el tiempo e invertir en procesos de innovación y cambio, precisamente por la visibilidad e impacto de lo desarrollado en *lo público*, han de hacer frente a una estrategia de cero errores. Lo realizado desde las administraciones ha de ser escrutado con precisión y ello en un contexto como el analítico, que emerge ahora y en el que existe siempre el riesgo sobre los resultados finales a obtener.

**Fortalezas.** Sin embargo, las Administraciones Públicas tienen importantes puntos de apoyo para el de-

sarrollo de estrategias de analítica basada en los datos. Desde luego el gran volumen de datos a los que se tiene acceso es sin duda un elemento fundamental. Pero no solo. La administración ha desarrollado un conjunto de mecanismos para controlar los procesos y los recursos. En un ámbito de reciente desarrollo y que toca aspectos centrales del derecho de los individuos a su intimidad, estos mecanismos de control se convierten en una fortaleza indispensable para desarrollar proyectos de analítica bajo una mirada atenta a los intereses de los diferentes grupos de interés. Por otro lado, tanto las limitaciones presupuestarias como la regulación de los procesos de gestión internos hace emerger una necesidad compartida de mejora: en muchos casos soportada en unos procesos eficientes en el desarrollo de nuevas capacidades en los profesionales. Por último no es desdeñable el contenido del trabajo del empleado público. La vocación de servicio a la comunidad es un elemento central en el significado atribuido al trabajo, y, en ello, las Administraciones Públicas tienen un fuerte aliado al posicionar al ciudadano como eje sobre el que desarrollar la gestión de recursos humanos.

**Amenazas.** El contexto en el que se desarrolla la analítica para recursos humanos es el típico en el que un importante número de actores confluyen y comienzan a desarrollar elementos en ausencia de fuertes sinergias y modelos claros de trabajo. La relevancia mediática del *Big Data* hace que en no pocas ocasiones se piense que por el solo hecho de tener grandes volúmenes de datos ya tenemos soluciones a muchos problemas. Hemos visto que no es así: la importancia de las preguntas a realizarse, del foco de los proyectos y del necesario desarrollo metodológico no debe menoscabarse. Por otro lado, y relacionado con la anterior visión, existe en el campo una importante ausencia de modelos y teorías sobre cómo desarrollar en la práctica la analítica. ¿Es algo novedoso que requiere de modelos específicos? ¿Hay evidencia empírica acerca de las bondades de esta forma de actuación? ¿Se conocen qué métodos y herramientas funcionan y cuáles no? Se requiere de avance en estas cuestiones y, como no, en el desarrollo de marcos de acuerdo y trabajo en el que los profesionales de recursos humanos vayan progresivamente estableciendo las posibilidades y casos de uso de esta aproximación. Por último una amenaza final tiene que ver con la capacitación de los profesionales de recursos humanos. Más allá de la utilización de las herramientas que provee la ofimática básica, en la moderna analítica de recursos humanos se requiere de una mayor capacitación en el terreno de lo científico, de la estadística y del manejo de datos por poner solo algunos ejemplos.

**Oportunidades.** Pero es verdad que el contexto actual es enormemente favorecedor para el desarrollo

de la analítica en las Administraciones Públicas. Un primer elemento de apoyo serán unas capacidades tecnológicas como antes no habíamos visto y que podrán ser adaptadas fácilmente a los requerimientos de la función pública. Por otro lado la eclosión de la actividad en el sector privado ha hecho que se desarrollen un importante número de casos de uso y buenas prácticas. La administración puede mirar hacia ellos para adaptarlos a sus necesidades teniendo un modelo básico de qué es lo que ha funcionado. Por último el Big Data en términos generales está encima de la mesa de las Administraciones Públicas y estas se están involucrando y cooperando a diferentes niveles para incorporar en sus procesos estos desarrollos tecnológicos. En su agenda habrá de estar también la analítica para recursos humanos.

En definitiva un panorama sumamente alentador pero que no está exento de amenazas y ante el que las Administraciones Públicas han de actuar apoyándose en sus fortalezas y limitando sus debilidades.

## 5. Conclusiones

El panorama descrito anteriormente nos lleva necesariamente a considerar un equipo de Recursos Humanos con un conjunto de competencias diferenciales respecto de las que hasta el momento imperaban. En un entorno de esta complejidad se comprende fácilmente cómo se requiere del profesional dedicado a la gestión de recursos humanos una capacitación diferencial, un desarrollo de la función más interdisciplinar y, en definitiva, una mirada estratégica diferente sobre el valor añadido de la gestión de recursos humanos. Así, sin duda parece necesaria la capacitación en competencias relacionadas: (a) con el uso de las redes sociales, los aspectos legales vinculados con la protección de la privacidad y el uso de la información que contempla la huella digital de los empleados y/o candidatos; (b) con un mayor conocimiento de los elementos económico-financieros que progresivamente se irán utilizando como criterio de valoración del capital humano de las organizaciones; y (c) con las habilidades básicas para el tratamiento y análisis de los datos relacionados con la fuerza laboral. La capacitación en estos aspectos permitirá que el profesional de los recursos humanos pueda hacer frente a un entorno de trabajo dominado por los equipos de trabajo y los proyectos interdisciplinares. Este es un aspecto fundamental que las Administraciones Públicas han de ir incorporando en sus planes de formación y desarrollo

de cara a que los profesionales de recursos humanos puedan incorporar progresivamente estas tecnologías.

Sumado a lo anterior es necesario considerar la cuestión de los parámetros legales y sociales asociados a la utilización de los datos que el trabajador genera en su quehacer diario. Desde luego la proyección de desarrollo de la analítica para recursos humanos está vinculada con un crecimiento de los datos (y los tipos de datos) que se utilizan. Ello implica que, como hemos visto, en un futuro próximo, en el que la Internet of Things nos permitirá monitorizar gran parte de nuestra actividad laboral, obtendremos ingentes cantidades de datos acerca de lo que un trabajador hace. ¿Qué normas sociales rigen la utilización de estos nuevos datos? ¿Cuáles son los usos permitidos? Todo ello nos lleva a plantearnos el derecho del trabajador a que los datos generados durante su trabajo (o incluso los generados fuera de su trabajo) puedan ser o no utilizados por su compañía para obtener ventaja competitiva. Sin duda esta cuestión tiene una enorme importancia y aún no tenemos un marco de referencia claro, al contrario, requiere de un conjunto de visiones compartidas y de reglas de comportamiento de las que nos hemos de dotar como profesionales.

Así las cosas, estos primeros años de desarrollo práctico de la analítica para recursos humanos permiten ofrecer ya algunas ideas claves en forma de lecciones aprendidas (Goldstein, Raza & Saraf, 2013): (a) no esperar a obtener datos perfectos, no existen; (b) empezar con proyectos piloto; (c) alinear el proyecto con la estrategia de negocio de la organización, en el caso de las Administraciones Públicas el servicio al ciudadano; (d) tener una sistemática a largo plazo; (e) desarrollar profesionales; y (f) no olvidar que los datos son sobre personas.

## Referencias

- Aguado, D. (2016). Redes Sociales en la Selección de Personal. Analizando sus propiedades psicométricas. *Observatorio de Recursos Humanos y Relaciones Laborales* 108, 48-53.
- Aguinis, H., & Bradley, K. J. (2015). The secret sauce for organizational success: managing and producing star performers. *Organizational Dynamics*, 44, 131-168.
- Bartram, D. (2009). The International Test Commission Guidelines on Computer-Based and Internet-Delivered Testing. *Industrial and Organizational Psychology* 2 (1), 11-13.
- Bersin, J. (2004). *The blended learning book: Best practices, proven methodologies, and lessons learned*. John Wiley & Sons.
- Boudreau, J. W., & Ramstad, P. M. (2002). Strategic HRM measurement in the 21st century: From justifying HR to strategic talent leadership. *Working Paper Series 02-15*, Center for Advanced Human Resource Studies (CAHRS).
- Boudreau, J. W., & Ramstad, P. M. (2007). *Beyond HR: The new science of human capital*. Harvard Business Press.
- Brown, H. (2010). For obese people, prejudice in plain sight. *New York Times*.
- Brown, V. R., & Vaughn, E. D. (2011). The writing on the (Facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business and Psychology*, 26, 219-225.
- Cheng, H., Zheng, N., & Qin, J. (2005). Pedestrian detection using sparse gabor filter and support vector machine. *IEEE Intelligent Vehicles Symposium*.
- Chien, C. F., & Chen, L. F. (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Systems with Applications*, (34), 280-290.
- Cho, V., & Ngai, E. W. T. (2003). Data mining for selection of insurance sales agents. *Expert Systems with Applications*, 20, 123-132.
- Delgado-Gómez, D., Aguado, D., López-Castromán, J., Santacruz, C., & Artés-Rodríguez, A. (2011). Improving sale performance prediction using support vector machines. *Expert Systems with Applications*, 38, 5129-5132.
- Delgado-Gómez, D., Carmona-Vázquez, C., Bayona, S., Ardoy-Cuadros, J., Aguado, D., Baca-García, E., Lopez-Castroman, E. (2015) Improving impulsivity assessment using movement recognition: A pilot study. *Behavioral Research Methods*, 47 (3).
- Downes, S. (2008). Places to go: Connectivism & connective knowledge. *Innovate: Journal of Online Education*, 5(1), 6.
- Fitz-Enz, J. (2010). *The New HR Analytics: Predicting the Economic Value of Your Company's Human Capital Investments*. AMACOM, a division of American Management Association. New York, USA.
- Goldstein, J.K., Raza, S., & Saraf, O. (2013). Lessons from Leaders. Sidebar in Brosnan, M. L., Farley, C. S., Gartside, D., & Tambe, H. (2013). *How well do you know your workforce? Outlook: Accenture's Journal of High-Performance Business*.
- Gosling, S. D., Augustine, A. A., Vazire, S., Holtzmann, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking*, 14, 483-488.
- Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1), 389-422.
- Harris, J. G., Craig, E., & Light, D. A. (2011). Talent and analytics: new approaches, higher ROI. *Journal of Business Strategy*, 32(6), 4-13.
- Haya, P. & García, P.E. (2018). El análisis de redes organizacionales: Concepto, metodología y aplicaciones. *Observatorio de Recursos Humanos y Relaciones Laborales* 131.

- Infoempleo-Adecco (2016). *Redes sociales y mercado de trabajo*. Disponible en <http://www.adeccorientaempleo.com/webwp/wp-content/uploads/2017/02/Informe-2017-Empleo-y-Redes-Infoempleo-Adecco-2.pdf> (descargado 17 Mayo de 2017).
- Kluemper, D. H. (2013). Social network screening: Pitfalls, possibilities, and parallels in employment selection. In Tanya Bondarouk, Miguel R. Olivas-Luján (Eds.) *Social Media in Human Resources Management (Advanced Series in Management, Volume 12)* (1-21). Emerald Group Publishing Limited.
- Lievens, F., & Chapman, D. S. (2009). Recruitment and selection. *Handbook of Human Resource Management*. 133-154. Sage Publications. California. USA.
- Markovikj, D., Gievska, S., Kosinski, M., & Stillwell, D. (2013). Mining facebook data for predictive personality modeling. In *Proceedings of the 7th international AAAI conference on Weblogs and Social Media (ICWSM 2013)*, Boston, MA, USA (pp. 23-26).
- Nikolaou (2014). Social networking web sites in job search and employee recruitment. *International Journal of Selection and Assessment* 22 (2), 179-189.
- Observatorio del Sector Público de Informática el Corte Inglés (2016). *Big Data en las Administraciones Públicas. Documento de Conclusiones*. Disponible en [http://www.administraciondigital2020.com/downloads/bigdata\\_ospi.pdf](http://www.administraciondigital2020.com/downloads/bigdata_ospi.pdf) (descargado 6 de febrero de 2018).
- Rubio, V.J., Aguado, D., Delgado-Gómez, D., Márquez, M.O., & Cernada, R. (2017). Can Big Data analysis be useful for assessing personality? A new approach to personality judgments based on movement recognition. Comunicación presentada en *14<sup>th</sup> European Conference on Psychological Assessment*. Lisbon, 5-8 July, 2017.
- Pavani, S., Delgado, D., & Frangi, A. (2010). Haar-like features with optimally weighted rectangles for rapid object detection. *Pattern Recognition*, 43(1), 160-172.
- Pin, J.R., García, P., & Peiró-Barra, A. (2012). *El papel estratégico de la dirección de personas en la Administración Pública española*. Cátedra José Felipe Bertrán de Gobierno y Liderazgo en la Administración Pública. IESE Business School. Universidad de Navarra. Disponible en <http://www.iese.edu/research/pdfs/ESTUDIO-261.pdf> (descargado 6 de febrero de 2018).
- Quercia, D., Kosinski, M., Stillwell, D., & Crocworth, J. (2011). Our twitter profiles, our selves: Predicting personality with twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, IEEE (pp. 180-185). IEEE.
- Quinn, A., Rycraft, J. R., & Schoech, D. (2002). Building a model to predict caseworker and supervisor turnover using a neural network and logistic regression. *Journal of Technology in Human Services*, 19 (4), 65-85.
- Reynolds, D. H., & Dickter, D. N. (2010). Technology and employee selection. *Handbook of employee selection*. 171-194. Taylor & Francis Group. New York, USA.
- Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, (40), 601-618.
- Roth, P. L., Bobko, P., Van Iddekinge, C. H., & Thatcher, J. B. (2016). Social media in employee-selection-related decisions a research agenda for uncharted territory. *Journal of Management* 42, 269-298.
- Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of applied psychology*, 87(1), 66-80.
- Salazar-Calle, J. & Romero-Sánchez, D. (2015). *Usa LinkedIn como si fueras un experto*. Auto-editado.
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M. & Ungar, L. H. (2013). *Personality, gender, and age in the language of social media: The open-vocabulary approach*. *PloS one*, 8(9), e73791.
- Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel methods for pattern analysis*. Cambridge: Cambridge university press.
- SIOP (2017). *Social Media and Selection: a brief history and practical recommendations*. SIOP White Paper Series. Disponible en: [http://www.siop.org/WhitePapers/Visibility/Social\\_Media\\_and\\_Selection\\_FINAL.pdf](http://www.siop.org/WhitePapers/Visibility/Social_Media_and_Selection_FINAL.pdf) (descargado 15 de marzo de 2018).
- Tay, F., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega*, 29(4), 309-317.
- Tippins, N. T. (2009). Internet alternatives to traditional Proctored Testing: Where Are We Now? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2 (1), 2-10.
- Tong, S., & Koller, D. (2002). Support vector machine active learning with applications to text classification. *The Journal of Machine Learning Research*, 2, 45-46.
- Tzeng, H. M., Hsieh, J.-G., & Lin, Y. L. (2004). Predicting nurses' intention to quit with a support vector machine: A new approach to set up an early warning mechanism in human resource management. *Computers, Informatics, Nursing*, 22, 232-242.
- Ulrich, D., & Dulebohn, J. H. (2015). Are we there yet? What's next for HR? *Human Resource Management Review* 25(2), 188-204.
- Yang, C. C., Lin, W. T., Chen, H. M., & Shi, Y. H. (2009). Improving scheduling of emergency physicians using data mining analysis. *Expert Systems with Applications*, 36, 3378-3387.
- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 1036-1040.